The effect of birthweight on childhood cognitive development in a middle-income country*

Florencia Torche† and Ghislaine Echevarría‡

* *Preliminary draft, please do not quote or cite without the authors' permission*. This research uses the restricted-access version of the SIMCE test scores dataset, produced by the Ministry of Education, Chile and of the Chilean birth registry dataset, produced by the Ministry of Health, Chile. The authors thank both Ministries for providing these datasets. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Ministries. The authors thank Francisco Mardones MD for help in locating the datasets, and Jere Behrman and Alejandro Corvalán for helpful comments.
† Department of Sociology, New York University. Email: florencia.torche@nyu.edu
‡ Faculty of Medicine, Universidad Católica de Chile. Email: gechevar@med.puc.cl

The effect of birthweight on childhood cognitive development in a middle-income country

Background: *Intra-uterine growth is a powerful predictor of infant mortality and of health, developmental, and socioeconomic outcomes in adulthood. The question about whether this relationship is causal rather than driven by unobserved characteristics of low-weight infants is, however, still open. We use twin models to examine the hypothesis that in-utero growth has a detrimental impact on cognitive development in childhood.*

Methods: *We merge birth registry information on birthweight with standardized Math and Spanish test scores for all 4th graders in Chile to create a prospective dataset. Twin fixed effects models are used to estimate the causal effect of intra-uterine growth on test scores. Fixed effect estimates are compared with traditional regression results in a cross-section of births to gauge the omitted variable bias emerging from unobserved genetic, maternal, and pregnancy-related factors in cross-sectional models.*

Results: *Birthweight differences within twin pairs have a substantial effect on test scores. A 400-gram increase in birthweight results in a 15% standard deviation-increase in Math scores. The effect is larger among (estimated) monozygotic than dizygotic pairs, reaching more than 20% standard deviation. The effect varies across family socioeconomic status: It is strong among disadvantaged families but it nearly disappears among advantaged ones.*

Conclusion: *Scarcity of uterine resources resulting in intra-uterine growth restriction affects cognitive development in early childhood. This effect interacts with parental behavior, which varies across socioeconomic status, so that low-SES families reinforce the effect of birthweight and high-SES families fully compensate for it. Findings are particularly relevant in the developing world, where intra-uterine growth restriction is the main determinant of low birthweight.*

Keywords: *Birthweight, twin studies, zygosity, fixed effects, cognitive development, developing countries.*

The effect of birthweight on cognitive development in a middle-income country.

Introduction. Birthweight is a powerful marker of early disadvantage. It is the main predictor of neonatal and infant mortality(Mathews and MacDorman 2008) and is related to health, development and wellbeing in childhood and adulthood(Barker, Gluckman, Godfrey, Harding, Owens, and Robinson 1993; Boardman, Powers, Padilla, and Hummer 2002; Brooks, Byrd, Weitzman, Auinger, and McBride 2001; Case, Fertig, and Paxson 2005; Currie and Hyson 1999; Hales, Barker, Clark, Cox, Fall, Osmond, and Winter 1991; Matte, Bresnahan, Begg, and Susser 2001; Pallotto and Kilbride 2006; Paneth 1995; Strauss 2000). To date, however, most research studying the influence of birthweight on later outcomes examines differences in a cross-section of births using regression models. Even if these models control for observed characteristics such as family socioeconomic status, maternal smoking, and pregnancy risk factors, they leave open the possibility of unobserved factors driving the influence of birthweight. If infants born low-weight differ in terms of genetic endowments, maternal behaviors, or any other factor unobserved by the researcher and consequential for later outcomes, failure to control for them will result in a spurious association. For example, the same genetic traits that determine low birthweight may affect individual educational and economic attainment, so that a policy increasing birthweight would have little effect on later achievements. To the extent that spuriousness cannot be ruled out, the policy implications of the findings are unclear.

A powerful strategy to address the unobserved selectivity of low birthweight is to examine variation in weight between twins through fixed effects models. This approach relates within twin-pair differences in birthweight with differences in the outcome of interest to identify a causal influence. Because twins share the uterine environment, the model accounts for confounders related to the mother and the pregnancy. Because twins have, by definition, the same gestational age, the observed effect of birthweight emerges entirely from intra-uterine growth, rather than from length of gestation. Recent studies use twins examine the effect of in-utero growth on perinatal and infant mortality(Almond, Chay, and Lee 2005; Conley, Strully, and Bennett 2006), early infancy development(Datar and Jacknowitz 2009), and adult outcomes such as education, BMI, height, and wages (Behrman and Rosenzweig 2004; Black, Devereux, and Salvanes 2007; Oreopoulos, Stabile, Walld, and Roos 2008; Royer 2009). Exploiting a twin fixed effects methodology, this paper addresses two questions: Does intra-uterine growth have an effect on children's cognitive development (measured by test scores)? And, does this effect vary across family socioeconomic circumstance?

The key assumption in the fixed effect approach is that weight differences within twin pairs emerge from random differences in access to nutritional intakes resulting, for example, from position in the uterus or umbilical cord insertion in the placenta. Research supports this assumption, showing that twins grow at the same rate as singletons up to the 30th week of gestation (Cleary-Goldman and D'Alton 2008). After that, the positive relationship between gestational age and fetal size cannot be maintained due to the restricted capability of the uterine environment to nurture two fetuses in a uniparious species, leading to weight differences between twins (Blickstein 2004; Blickstein 2005b).

Although all twins face scarcity in the womb, the magnitude and etiology of intra-uterine growth differences depend on zygoticity. Dizygotic twins (DZ) result from the separate fertilization of two ova, and share, on average, 50% of their genes, the same as regular siblings. Monozygotic twins are formed from the division of a single ovum after fertilization. Even though they are not genetically identical (Gringras and Chen 2001; Hall 2003; Machin 2009), monozygotic twins share most of their genes. Zygosity affects also the access to uterine resources. Dizygotic twins virtually always have separate placentas with no obvious vascular connections (Hall 2003). In contrast, most monozygotic twins share a single placenta, which places them at heightened competition for uterine resources (Hall 2003; Hanley, Ananth, Shen-Schwarz, Smulian, Lai, and Vintzileos 2002; Nikkels, Hack, and van Gemert 2008; Sakata, Utsu, and Maeda 2006). While access to nutrients may differ between dizygotic twins because of differences in the mass and function of the placenta or placental lesions, these differences are usually moderate (Bajoria, Sooranna, Ward, and Hancock 2002; Blickstein 2005a; Victoria, Mora, and Arias 2001). In contrast, among monozygotic twins sharing a placenta, severe differences in nutritional intake may emerge from differential umbilical cord insertion and uneven placental sharing (De Paepe, Shapiro, Young, and Luks 2010; Hall 2003). Weight discrepancies are thus larger, and the lighter monozygotic twin is likely to be genuinely growth restricted (Blickstein 2005a; Hack, Derks, Elias, Franx, Roos, Voerman, Bode, Koopman-Esseboom, and Visser 2008; Sherer and Divon 1997). As a result, monozygotic twinning controls not only for mother- and pregnancy-related confounders but also for genetic variation, and offers a setting characterized by random –and acute– environmental determinants of intra-uterine growth.

To date, studies using twin models have focused on a handful of industrialized countries. Researchers have questioned the generalizability to other national contexts, but empirical evidence is lacking. In addition, research has focused on early outcomes such as infant mortality and morbidity; or on outcomes in adulthood such as completed education and wages. Less evidence exists on intermediate outcomes such as cognitive development in childhood, which likely connect early developments with adult attainment. Furthermore, researchers usually compare fixed effect with cross-sectional regression estimates to gauge unobserved selectivity bias in cross-sectional models. But because twins share the same gestational age, meaningful comparison requires controlling for gestational age in cross-sectional samples, so that estimates capture solely the effect of intra-uterine growth in both cases.

This study contributes to the literature in three respects. First, we use twin fixed effect models to examine the effect of intra-uterine growth on Math and Spanish test scores of Chilean 4th graders in 2008. Fixed effect models explicitly account for unobserved genetic, maternal, and pregnancy-related characteristics related to birthweight. Test scores are an important outcome because they measure early cognitive development, and predict later educational attainment and labor market outcomes(Currie and Thomas 2001; Rose 2006), providing an intermediate variable linking short- and long-term outcomes of birthweight. Second, a middle-income country with high economic inequality, Chile offers a useful setting to test the generalizability of the effect of birthweight. Third, by using an adjusted measure of intra-uterine growth orthogonal to gestational age, we provide valid comparisons between fixed effects and cross-sectional estimates of the effect of in-utero growth.

Methods:

Data: We construct a longitudinal dataset by merging standardized test scores with birth registry information on birthweight for all 4th graders in 2008. Test scores were obtained from the System of Measurement of Educational Quality (Sistema de Medición de la Calidad de la Educación, or SIMCE in Spanish) dataset. SIMCE is a standardized pencil-and-paper test administered to all 4th graders in the country in November 2008 by the Chilean Ministry of Education. The test includes multiple choice and short answers, and is given to students in their classrooms. Scores from the

Mathematics and Spanish were used for this analysis. Birthweights were obtained from the Chilean Birth Registry database established by the Chilean Ministry of Health, which records all births in the country. The database is constructed on the basis of live-birth certificates, a form completed by the professional (MD or certified midwife) who attends the delivery (99.8% of deliveries are attended by a professional in Chile), and measured with little error (Mardones, Villarroel, Karzulovic, Barja, Arnaiz, Taibo, and Mardones-Restat 2008). The birth certificate contains information on birthweight, gestational age, maternal residence, and demographic characteristics. The restricted-access version of both data sources contains an individual identifying number assigned to all Chilean citizens at time of birth. We used identifying numbers to merge these sources. Given that the normative age to enter 4th grade is 9 years old in Chile, most of our merged sample was born in 1998 (59.3%) and 1999 (35.0%). A small percent were born in 1997 (5.8%) and 2000 (0.01%). The resulting merged dataset has 220,062 valid observations. From this dataset, we identified 2,474 twins with information on all relevant variables. The use of administrative data has important advantages. First, it provides a census of 4th graders in 2008 with no sample error. Second, it avoids the potentially severe recall bias and measurement error affecting self-reported birthweight information. Zygosity information is not available in the dataset, so we estimate the number of monozygotic twin pairs as explained in detail in the next subsection.

Variables: Our measure of intra-uterine growth is the standardized value of gestational age-specific birthweight, obtained as the difference between individual birthweight and gestational week-specific mean birthweight, and dividing by the gestational week-specific standard deviation of birthweight. This measure is similar to the standard definition of "small for gestational age" –

birthweight less than the 10th percentile for gestational week(Resnik 2002)– but it captures variation throughout the entire distribution rather than establishing a single threshold. Given sex differences in fetal growth, standardized measures were separately obtained for male and female fetuses. Our standardized measure is orthogonal to length of gestation (with a correlation of -.007 in our sample). The standard deviation of birthweight is approximately 400 grams throughout the distribution, so a 1-unit increase in the standardized measure of intra-uterine growth involves a gain in birthweight of about 400 grams. Values of mean and standard deviations are obtained from the entire dataset of births, including twins and singletons. Because twins are usually lighter than singletons of comparable gestational age, they are overrepresented in the lower end of the distribution.

The dependent variables are Math and Spanish test scores. These variables have, by construction, a mean of 250 and a standard deviation of 50. All models include the following controls: Maternal age (less than 19 years of age, 20-29, 30-34, 35 or more), education (primary, secondary, post-secondary), urban residence (urban=1), employment status (mother works at the time of delivery=1), and infant's sex (male=1).

Analysis. We obtain twin fixed effect estimates of the effect of birthweight on test scores, and compare them to regression estimates of the cross-sectional sample of both, singleton and twin births. The fixed-effects model is formulated as follows:

$$t_{ijk} = \beta_0 + bw_{ijk}\beta_1 + s_{ijk}\beta_2 + X_{jk}'\beta_3 + A_{jk} + \varepsilon_{ijk}$$

Where t identifies the dependent variable, test scores, of child *i* born to mother *j* in birth *k*. bw refers to birthweight, s identifies the sex of the child, X_{jk} is a vector of mother- and pregnancy-specific characteristics such as maternal age, rural residence, and education, A is a vector of mother- and birth-specific unobserved factors (for example, genetic factors or pregnancy behaviors), and ε is an idiosyncratic error term. Cross-sectional estimates may be biased because no controls are included for factors in A, if these factors are correlated with both birthweight and test scores. One strategy to control for factors included in A is to differentiate the data within twin pairs. For a pair formed by twins 1 and 2, we obtain the differenced equation:

$$t_{2jk} - t_{1jk} = (\beta_{02}-\beta_{01}) + \beta_1(bw_{2jk}-bw_{1jk}) + \beta_2(s_{2jk}-s_{1jk}) + (\varepsilon_{2jk}-\varepsilon_{1jk})$$

Because the unobserved mother- and pregnancy-specific factors included in A do not vary across twins, the term a drops from the differenced equation. Given that we have information on child's sex, and under the assumption that within-twin differences in ε_{ijk} are independent from within-twin differences in bw_{ijk}, β_1 is consistent. This assumption is more likely to hold among monozygotic than dizygotic twins because the former share most of their genes. We estimate models for the entire twin sample, for the same-sex and mixed-sex subsamples, and for the group estimated to be monozygotic. Given that monozygotic twins share most their genes and that they face stronger competition for uterine resources than their dizygotic counterparts, we expect differences in birthweight within monozygotic pairs to be a much better proxy for individual twin levels of nutrition than differences within dizygotic pairs.

Lacking information about zygosity, we use a method proposed by Weinberg(Weinberg 1902) to obtain an estimate of the number of monozygotic pairs. This method is based on the fact that mixed-sex twins are necessarily dizygotic, while same-sex twins may be either dizygotic or monozygotic. Under the assumption that the probability of a male birth equals the probability of a female birth and that the sexes in a dizygotic pair are independent, the rate of dizygotic twinning is twice the rate of mixed-sex births. The monozygotic twinning rate is the difference between the rates of same-sex and mixed-sex twin pairs. The Weinberg method makes some questionable assumptions, including the rare occurrence of monozygotic twins of unlike sex, the fact that when monozygotic twins are obviously discordant they are judged dizygotic, and the suggestion that the actual frequency of same sex twins is higher than expected in dizygotic pairs (Hall 1996; Hall 2003). However, it produces extremely accurate estimates (Fellman and Eriksson 2006). Naturally, it is not possible to identify specific twin pairs by zygosity, but a weighted average strategy can be used to estimate birthweight effects by zygosity, as implemented by Conley et al. (Conley, Strully, and Bennett 2006). In a final step, we examine variation in the effect of birthweight across family's socioeconomic resources, using mother's education as a proxy.

Results.

Descriptive Analysis. Table 1 presents the descriptive statistics of all variables in the twin and singleton samples. As expected, birthweight, intra-uterine growth, and gestational age in twins is significantly less than singletons. Singletons perform better, on average, in Math and Spanish tests. Women giving birth to twins are, on average, more socioeconomically advantaged, but the differences are small. The secondary sex ratio is lower among twins (47.3% vs. 50.3% for singletons) probably related to a well-documented excess of females in spontaneous monozygotic twins (Hall 2003).

Table 1 about here

An important requirement of twin models is that there is sufficient within twin-pair variation in birthweight to identify its effect. Figure 1 displays the frequency distribution of differences among same-sex and mixed-sex twin pairs, and shows substantial variation. The average difference within twin pairs is 283 grams, with about half of the pairs having differences of 250 grams or more, and about 10% displaying differences of 600 grams or more.

Figure 1 about here

The effect of intra-uterine growth on test scores. Table 2 examines the effect of intra-uterine growth on Math and Spanish test scores using linear regression models on cross-sectional samples of singletons and twins; and fixed effect models in the twin sample. Models include controls for sex, maternal age, education, urban status, and employment status at time of birth. All cross-sectional specifications adjust for the clustering of standard errors within twin pairs in the calculation of 95% confidence intervals and p-values. Cross-sectional results for Math test scores indicate that a one-unit increase in fetal size is associated with 2.2 [CI: 2.01, 2.46] points increase in test scores among singletons, and 2.6 [CI: 0.26, 4.84] points among twins, net of socio-demographic covariates. Given that one unit of the independent variable is equivalent to approximately 400 grams, and that test scores have a standard deviation of 50 by construction, a 400-gram increase would result in test score gain of about 5% (2.5/50) of a standard deviation. Comparable figures for Spanish test scores are 1.2 [CI: 0.94, 1.38] and 2.6 [CI: 0.37, 4.80], corresponding to 2.4% and 5.2% of a standard deviation, respectively. The cross-sectional effect of intra-uterine growth on test scores is substantially very small.

Table 2 about here

When twin fixed-effects are used to control for unobserved characteristics of the child, the effect increases substantially, reaching 15% standard deviation for Math and 11% for Spanish (7.41/50 and 5.31/50, respectively). The larger fixed effects estimate has two alternative interpretations. First, it suggests that the scarcity of nutrients and oxygen resulting from uterine competition, rather than unobserved factors, has a substantial effect on test scores. Second, intra-uterine growth may be negatively correlated with endowments determining test scores, or nutrient intakes may be provided less to fetuses with higher endowments in the cross-section. To the best of our knowledge, no empirical evidence suggest the latter alternative, therefore we interpret this finding as evidence for a detrimental effect of resource scarcity in utero.

Exploiting Zygosity to Examine Effect of Uterine Environment: As a next step, we use sex differences within twin pairs to explore the factors driving the effect of in-utero growth. In our sample, 75% of twin pairs are of the same sex and the 25% remaining are mixed-sex. Because monozygotic twins are virtually always of the same sex, they should be overrepresented in the same-sex subsample. If the effect of birthweight on test scores is driven by environmental factors in-utero, rather than unobserved genetic variation at the individual level, we expect the effect to be greater among same-sex twins.

Table 3 about here

This hypothesis is supported in table 3, which shows a larger effect of intra-uterine growth among same-sex twins. A 400-gram increase results in a gain of 8.7 [CI: 5.32, 12.12] points in Math test scores, and 6.9 [CI: 3.33, 10.53] points in Spanish, equivalent to 17% and 14% standard deviation, respectively. The effect is smaller among mixed-sex twin pairs. Because mixed-sex pairs are necessarily dizygotic, and dizygotic twins are much less subject to competition for uterine resources than dizygotic pairs (but are genetically different), this finding suggests that the factor accounting for the effect of intra-uterine growth on test scores is access to nutrition in-utero rather than genetic factors. These estimates, in association with the Weinberg rule, allow us to gauge the effect among monozygotic twins. According to Weinberg's formulation, the parameter estimate for MZ twins can be calculated as follows:

$$b_{MZ} = \frac{b_{SS} - b_{MS}\,(1-p_{MZ})}{p_{MZ}}$$

where b_{MZ} is the estimate among monozygotic twins, b_{SS} and b_{MS} are, respectively, the estimates for same-sex and mixed-sex pairs presented in table 3, and p_{MZ} is the proportion of monozygotic twins in the same-sex group. 602 twins in the sample belong to mixed-sex pairs and 1,872, to same-sex pairs, for a total sample of 2,474. According to Weinberg's rule, 602 twins in the same-sex group are dizygotic, and the remaining 1,270 are monozygotic, yielding a p_{MZ} value of .678 According to these calculations nearly half of the twin pairs (1,270/2,474) are monozygotic. This percentage is higher than reported for some regions of the world(Imaizumi 2003), but very similar to previous findings for Chile (Ivanovic, Llop, Alvear, Perez, Diaz, Leyton, Almagia, Larrain, Alvarez, Herrera, and Hazbun 2006). Using the parameter estimates from table 3 we obtain b_{MZ}=[8.721 –3.721 (.322)]/.678 = 11.10 for Math test scores and b_{MZ}=[6.929 –0.754 (.322)]/ .678 = 9.86 for Spanish (Calculation of standard errors, 95% CI, and P-values for these parameter estimates are provided in Supplementary data; available at IJE online). Thus 400-gram increase in birthweight results in a very large increase of 22% (11.10/50) standard deviation in Math test scores and 20% (9.86/50) standard deviation in Spanish test scores among monozygotic twins. This result supports the assertion that the main factor accounting for the beneficial effect of fetal growth on test scores is the availability of uterine resources rather than unobserved genetic attributes.

Socioeconomic Heterogeneity in the Effect of Intra-Uterine Growth: So far we have assumed that the effect of in-utero growth is homogeneous across the population. However, the homogeneity assumption is likely not satisfied. In particular, the literature on intra-household resource allocation suggests that parents may invest differentially depending on offspring's birthweight. Parents may reinforce initial disadvantage by allocating more resources to the high-weight twin or compensate by investing more in the low-weight twin. Theoretical models suggest parental behavior depends on the degree of equal concern across children, preference for equity versus productivity, distribution of endowments across children, and characteristics of the offspring's outcome (test scores, in this case) function (Becker and Tomes 1979; Behrman, Pollak, and Taubman 1982). Empirical evidence mostly favors reinforcement in the US (Behrman, Pollak, and Taubman 1982; Behrman, Rosenzweig, and Taubman 1994; Datar, Kilburn, and Loughran 2010; Griliches 1979) as well as developing contexts (Pitt, Rosenzweig, and Hassan 1990; Rosenzweig and Schultz 1982; Rosenzweig and Wolpin 1988; Rosenzweig and Zhang 2009).

Parental behaviors are not directly observable, but it is plausible that the extent of compensation varies according to the socioeconomic resources of the family. Several studies in industrialized world have considered this possibility but have found no substantial variation across socioeconomic status (Black, Devereux, and Salvanes 2007; Datar, Kilburn, and Loughran 2010). Variation may be more pronounced in an unequal middle-income country, however. Given lower levels of income and higher levels of inequality in Chile, we hypothesize that resource-constrained families may be more likely to target resources on the twin with higher chances of succeeding; while less-constrained families may be more able to compensate for observed disadvantages.

Table 4 about here

In order to assess heterogeneity in the effect of birthweight, table 4 presents a model with interactions between intra-uterine growth and our best proxy for family resources –mother's education– distinguishing primary, secondary, and post-secondary levels of schooling. The variation across maternal education is substantial. For Math test scores, a 400-grams increase in birthweight raises scores by about 12.4 [CI: 6.210, 18.50] points, i.e. by about 25% of a standard deviation (12.4/50) among low-education mothers. The effect declines to only 14% for mothers with secondary schooling, and a negligible 3.7% among mothers with post-secondary schooling. The pattern is the same for Spanish scores although effects are somewhat weaker, and is replicated also across same-sex pairs. This result is consistent with recent findings in behavioral genetics showing that the effect of shared environment on twins' IQ is much stronger in disadvantaged families than in affluent ones {Turkheimer, 2003 #496}. The variation in the effect of intra-uterine growth by maternal education suggests reinforcing behavior among poor families and compensatory behavior among better-off families –to the extent that the disadvantage associated with low intra-uterine growth fully disappears among advantaged families.

Discussion. Although the amount of variation explained by our models is small, our analysis indicates that intra-uterine growth has a substantial effect on children's cognitive development, as measured by test scores in primary school. The observed effect is larger among twin pairs estimated to be monozygotic than among those estimated to be dizygotic. Combined, these results suggest that the factor driving the effect of intra-uterine growth on cognitive development is the scarcity of environmental resources in utero, rather than unobserved confounders, including genetic ones. These findings are particularly relevant for developing countries, not only because they were obtained in Chile, but also because the source of variation in birthweight analyzed –intra-uterine growth– is the main determinant of low birthweight in the developing world. In fact, intra-uterine growth restriction emerging from factors such as deficient maternal nutrition, low pre-pregnancy BMI, short stature, and infectious diseases such as malaria affects up to one-third of births in poor countries (Bale, Stoll, Lucas, and Institute of Medicine (U.S.). Committee on Improving Birth Outcomes. 2003; Kramer and Victora 2001). Because these factors are likely to result in the type of uterine scarcity acutely experienced by monozygotic twins, our findings are potentially relevant for most of the low-weight infants in less advantaged regions of the world.

But twins are only about 1.5% of all births. Are these findings generalizable to the larger population of singletons? Twins are on average much smaller than singletons at birth. In our sample, for instance, singletons weight on average 3,373 grams while twins weight only 2,500 grams. The gap itself is not an issue as long as the factors accounting for the smallness of twins do not reflect omitted variables beyond the scarcity of uterine space and nutrients in a uniparious species (Blickstein and Kalish 2003). In other words, we rely on the assumption that the smallness of twins at birth should be "the result of twinning rather than the selectivity of parents who produce twins" (Behrman and Rosenzweig 2004: see pp. 591). While the literature supports this assumption, this may not be the case for singletons. Given that singleton pregnancies are not constrained by uterine space or nutrient scarcity, is a singleton weighting 2,500 grams comparable to a twin of the same weight? Research suggests that in-utero growth restriction among singletons may be related to congenital malformations, chromosomal anomalies, or inborn errors in metabolism (de Bie, Oostrom, and Delemarre-van de Waal 2010; Resnik 2002). To the extent that these negative factors are correlated with test scores, findings for twins will not generalizable to singletons of comparable birthweight. Furthermore, given that twins are at much higher risk of perinatal and infant mortality (Nikkels, Hack, and van Gemert 2008) than singletons, surviving twins may be positively selected in factors affecting test scores, further reducing comparability. Negative selectivity of twins is also possible. Even if "the smallness of twins result from twinning", twins may still be selected on endowments related to cognitive outcomes such as test scores. Addressing this question requires examination of the causes of twinning. Research identifies genetic factors, maternal age, race, height, BMI, fertility enhancing drugs, and folic acid supplementation as causes of DZ twinning; and problems during transport through the fallopian tube, conception in close proximity to oral contraception, and assisted reproductive technology as causes of MZ twinning (Hankins and Saade 2005; Hoekstra, Zhao, Lambalk, Willemsen, Martin, Boomsma, and Montgomery 2008; Scott 2002).

Table 5 about here

To assess the comparability of twins and singletons of similar birthweight and gestational age, we create a pooled sample over an area of common support defined in terms of birth outcomes and maternal characteristics by means of propensity score weighting. This strategy weights the singleton dataset to ensure balance with the twin dataset according to the following variables: Birthweight, gestational age, fetal growth, maternal age, education, urban status, and work status, and sex. Using this matched dataset, we predict Math and Spanish test scores using an indicator for twin. If singletons comparable to twins are negatively selected on factors affecting test scores, we should observe that being a twin has a positive effect on test scores. Findings, reported in table 5, reject this hypothesis. They show no significant differences in test scores between twins and comparable singletons, suggesting that growth-restricted singletons are not negatively (or positively) selected, and that findings from this study could potentially be extended to the larger singleton population affected by intra-uterine growth restriction.

Another concern refers to the determinants of in-utero growth differences within twin pairs, and whether they are similar to cross-sectional differences across singletons. We have argued that these factors refer to nutrient scarcity in-utero, which is more acute among monozygotic than dizygotic twins. However, in some cases of monozygotic twinning, heavier birthweight is not a marker of advantage, because the twin-to-twin vascular exchange results in poor outcomes for both, the lighter and the heavier twin. This is the case, in particular, with the twin-to-twin transfusion syndrome, which affects approximately 15% of monozygotic twins that share a placenta (Sakata, Utsu, and Maeda 2006). To the extent that twin-to-twin exchange results in detrimental outcomes for the heavier twin too, the observed effect will be attenuated and our estimates will underestimate the true effect of uterine growth.

In sum, our analysis indicates that in-utero growth has a substantial impact on cognitive development during childhood among twins, and probably among singletons. Given that intra-uterine growth is the main determinant of low birthweight in the developing world, the analysis is particularly relevant to less-advantaged regions of the world. The variation in the effect across levels of maternal education suggests that parental behavior may reinforce the influence of birthweight on educational achievement, increasing the vulnerability of the low-birthweight children among disadvantaged families; while high-resource families may be able to fully compensate for early disadvantage. More research accounting for heterogeneity of birthweight in diverse national contexts is necessary to examine variation in family compensatory behavior across socioeconomic status, and to further understand the influence of the "first injustice" over the life-cycle.

Table 1. Means (standard deviations) and percent distributions by type of birth, and comparison between twins and singletons means/proportions Chile 1997-1999.

	Twins	Singletons	p-value difference means/ proportions
Birthweight	2,500.21 (487.62)	3,373.38 (499.42)	<.001
Weeks of gestation	36.29 (2.09)	38.92 (1.47)	<.001
Intra-Uterine Growth	-0.677 (.883)	0.028 (.987)	<.001
Small for gestational age*	22.89% (.42)	5.30% (.22)	<.001
Math Test Scores	242.79 (56.2)	248.6 (54.8)	<.001
Spanish Test Scores	257.1 (53.3)	261.7 (53.4)	<.001
Age at test	10.07 (.433)	10.12 (.466)	<.001
Mother's Age			<.001
Less than 19 (1)	8.2%	16.1%	
19 – 29 (2)	46.1%	49.9%	
30 – 34 (3)	26.0%	20.4%	
35 – 39 (4)	16.4%	11.0%	
40 or more (5)	3.3%	2.7%	
Mother's education			<.001
Primary (1)	24.5%	26.5%	
Secondary (2)	55.3%	58.6%	
Tertiary (3)	20.2%	14.9%	
Urban residence	88.9%	87.6%	.018
Mother works	32.4%	26.8%	<.001
Male	47.3%	50.3%	.001
N	2,474	217,588	

* Below the 10^{th} percentile of gestational age-specific birthweight

Figure 1.Absolute value of birthweight difference among same-sex and mixed-sex twins. Twins born in Chile 1997-1999 and who took 4th grade standardized test in 2008.

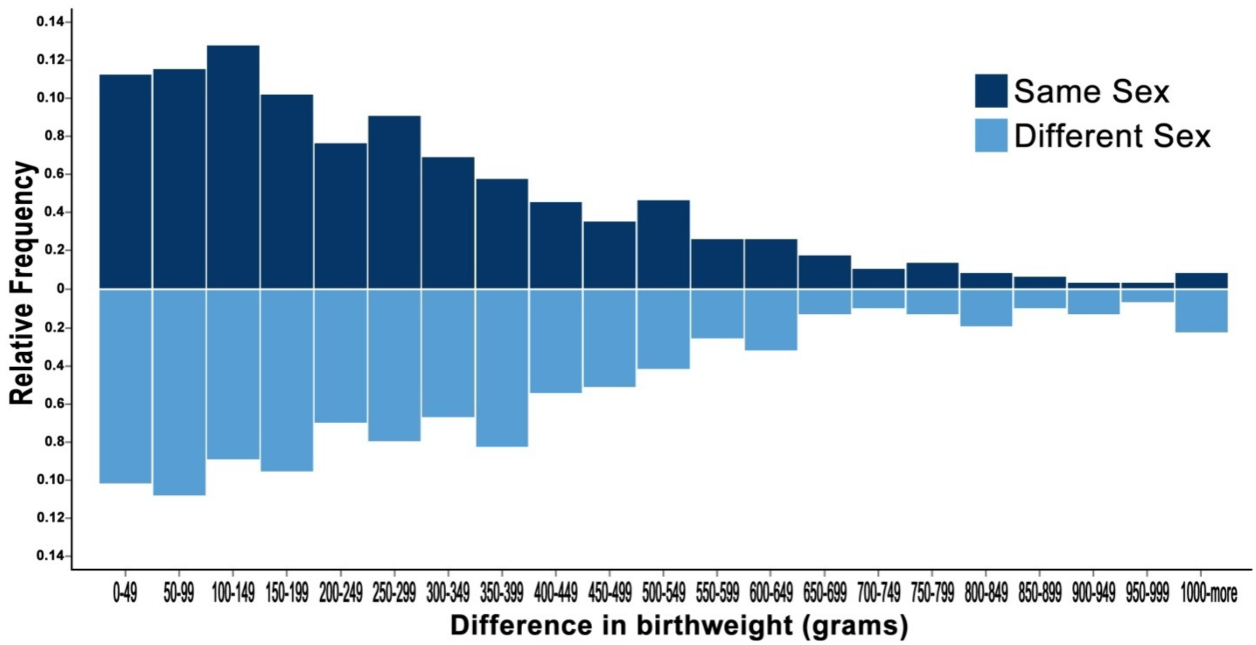

Table 2. Effect of intra-uterine growth on Math and Spanish test scores, OLS model in singleton and twin databases, and twin fixed-effects model. Children born in Chile in 1997-1999 and who took 4th grade standardized test in 2008 (95% confidence intervals in parentheses)[1].

	Math Test Scores									Spanish Test Scores								
	Singletons OLS			Twins OLS			Twins FE			Singletons OLS			Twins OLS			Twins FE		
	b	95% CI	*P*-value	b	95% CI	*P*-value	b	95% CI	*P*-value	b	95% CI	*P*-value	b	95% CI	*P*-value	b	95% CI	*P*-value
Intra-uterine Growth	2.23	(2.01, 2.46)	<0.001	2.55	(0.26, 4.84)	0.029	7.41	(4.33, 10.48)	<0.001	1.16	(0.94, 1.38)	<0.001	2.58	(0.37, 4.80)	0.022	5.31	(2.12, 8.50)	0.001
Mother's second. ed.[2]	16.48	(15.93, 17.02)	<0.001	17.96	(12.48, 23.43)	<0.001				13.99	(13.46, 14.52)	<0.001	17.20	(11.90, 22.49)	<0.001			
Mother's tertiary ed.[2]	42.46	(41.66, 43.26)	<0.001	53.63	(46.54, 60.72)	<0.001				35.24	(34.45, 36.02)	<0.001	42.78	(35.93, 49.64)	<0.001			
Mother's age 20-29[3]	0.96	(0.32, 1.62)	0.003	0.69	(-7.26, 8.63)	0.866				1.10	(0.46, 1.749)	0.001	0.83	(-6.83, 8.52)	0.832			
Mother's age 30-34[3]	3.58	(2.82, 4.35)	<0.001	3.84	(-4.61, 12.29)	0.372				3.68	(2.93, 4.43)	<0.001	5.77	(-2.41, 13.94)	0.167			
Mother's age 35-39	3.99	(3.11, 4.88)	<0.001	8.22	(-0.70, 17.13)	0.071				4.79	(3.92, 5.65)	<0.001	6.68	(-1.94, 15.30)	0.129			
Mother's age 40+	4.09	(2.62, 5.56)	<0.001	5.69	(-7.59, 18.98)	0.401				5.04	(3.59, 6.48)	<0.001	11.10	(-1.75, 23.95)	0.090			
Urban	1.89	(1.20, 2.59)	<0.001	-1.29	(-8.26, 5.68)	0.717				-0.96	(-1.64, -0.27)	0.006	-7.23	(-13.98, -0.49)	0.036			
Mother works	8.76	(8.20, 9.31)	<0.001	9.78	(4.89, 14.68)	<0.001				7.35	(6.808, 7.89)	<0.001	8.32	(3.59, 13.05)	0.001			
Male	4.30	(3.86, 4.74)	<0.001	4.74	(0.65, 8.83)	0.023	9.83	(4.46, 15.20)	<0.001	-12.40	(-12.83, -11.97)	<0.001	-12.38	(-16.33, -8.42)	<0.001	-6.27	(-11.84, -0.70)	0.027
Constant	224.6	(223.7, 225.4)	<0.001	219.1	(209.2, 229.0)	<0.001	247.0	(243.4, 250.5)	<0.001	251.4	(250.6, 252.2)	<0.001	249.8	(240.3, 259.4)	<0.001	267.7	(264.0, 271.4)	<0.001
Observations		217,588			2,474 (1237 pairs)			2,474 (1237 pairs)			217,588			2,474 (1237 pairs)			2,474 (1237 pairs)	
R-squared		0.084			0.146			0.028			0.072			0.112			0.012	

[1] Confidence intervals and *P*-values in OLS models are adjusted for clustering of standard errors at the mother's level.
[2] Reference group for mother's education is primary education
[3] Reference group for mother's age is less than 20 years old.

Table 3. Effect of intra-uterine growth on Math and Spanish test scores. Twin fixed effects for all twins, same-sex, and mixed-sex twin pairs. Children born in Chile in 1997-1999 and who took 4th grade standardized test in 2008 (95% confidence intervals in parentheses)[1].

	Math Test Scores									Spanish Test Scores								
	All Twins			Mixed Sex Twins			Same Sex Twins			All Twins			Mixed Sex Twins			Same Sex Twins		
	b	95% CI	*P*-value	b	95% CI	*P*-value	b	95% CI	*P*-value	b	95% CI	*P*-value	b	95% CI	*P*-value	b	95% CI	*P*-value
Intra-uterine Growth	7.41	(4.33, 10.48)	<0.001	3.72	(-3.14, 10.60)	0.287	8.72	(5.32, 12.12)	<0.001	5.31	(2.12, 8.49)	0.001	0.754	(-6.03, 7.54)	0.827	6.93	(3.33, 10.53)	<0.001
Male	9.83	(4.46, 15.20)	<0.001	9.99	(3.84, 16.14)	0.002				-6.27	(-11.84, -0.70)	0.027	-6.075	(-12.16, 0.01)	0.051			
Constant	247.0	(243.4, 250.5)	<0.001	248.2	(242.0, 254.3)	<0.001	251.2	(248.4, 253.9)	<0.001	267.7	(264.0, 271.4)	<0.001	268.4	(262.3, 274.5)	<0.001	264.8	(261.9, 267.7)	<0.001
Observations	2,474 (1237 pairs)			602 (301 pairs)			1872 (936 pairs)			2,474 (1237 pairs)			602 (301 pairs)			1872 (936 pairs)		
R-squared	0.028			0.038			0.026			0.012			0.013			0.015		

[1] Controls for maternal age, maternal education, maternal employment status, and urban residence dropped because variables do not vary within mothers.

Table 4. Effect of intra-uterine growth (IG) on Math and Spanish test scores across levels of mother's education. Twin fixed effects for all twins, same-sex, and mixed-sex twin pairs. Children born in Chile in 1997-1999 and who took 4th grade standardized test in 2008 (95% confidence intervals in parentheses).

	Math Test Scores									Spanish Test Scores								
	All Twins			Different Sex			Same Sex			All Twins			Different Sex			Same Sex		
	b	95% CI	*P*-value	b	95% CI	*P*-value	b	95% CI	*P*-value	b	95% CI	*P*-value	b	95% CI	*P*-value	b	95% CI	*P*-value
IG * Primary education	12.35	(6.20, 18.50)	<0.001	9.85	(-6.02, 25.72)	0.223	12.97	(6.45, 19.49)	<0.001	5.73	(-0.65, 12.12)	0.078	3.05	(-12.64, 18.74)	0.702	6.41	(-0.50, 13.31)	0.069
IG * Secondary education	6.82	(2.47, 11.17)	0.002	1.92	(-7.51, 11.35)	0.689	8.73	(3.85, 13.60)	<0.001	6.17	(1.66, 10.69)	0.007	-2.96	(-12.28, 6.37)	0.533	9.73	(4.57, 14.89)	<0.001
IG * Tertiary education	3.68	(-2.44, 9.80)	0.239	3.06	(-9.91, 16.03)	0.643	3.93	(-2.98, 10.85)	0.265	3.16	(-3.20, 9.52)	0.330	6.26	(-6.56, 19.08)	0.337	1.88	(-5.45, 9.20)	0.615
Male	9.82	(4.449, 15.20)	<0.001	9.83	(3.64, 16.02)	0.002				-6.19	(-11.77, -0.61)	0.030	-6.38	(-12.50, -0.27)	0.041			
Constant	247.0	(243.4, 250.5)	<0.001	248.3	(242.2, 254.5)	<0.001	251.1	(248.4, 253.9)	<0.001	267.7	(264.1, 271.4)	<0.001	268.4	(262.2, 274.5)	<0.001	265.0	(262.1, 268.0)	<0.001
Observations		2,474 (1237 pairs)			602 (301 pairs)			1872 (936 pairs)			2,474 (1237 pairs)			602 (301 pairs)			1872 (936 pairs)	
R-squared		0.032			0.040			0.030			0.013			0.017			0.018	

Table 5. Propensity score matching estimates of the effect of being a singleton (vs a twin) on Math and Spanish test scores[1].

	Math Test Scores			Spanish Test Scores		
	b	95% CI	*P*-value	b	95% CI	*P*-value
Twin	-0.62	(-3.16, 1.92)	0.633	-1.72	(-4.14, 0.71)	0.165
Constant	247.2	(245.9, 248.4)	<0.001	263.0	(261.7, 264.2)	<0.001
Observations		220,062			220,062	

[1] Propensity score matching produces a matched sample of twins and singletons, in which singletons are similar to twins in terms of birthweight, gestational age, uterine growth, sex, maternal age, education, and work status. Matching method is nearest neighbor with replacement and five neighbors in the singleton sample for each twin observation. Alternative methods to calculate matched outcome included single nearest neighbor with and without replacement, kernel matching with Epanechnikov kernel, and local linear regression matching provide similar results. Estimates obtained using routine developed by Leuven and Sianesi(Leuven and Sianesi 2003).

References

Almond, D., K. Y. Chay, and D. S. Lee. 2005. "The costs of low birth weight." *Quarterly Journal of Economics* 120:1031-1083.

Bajoria, R., S. R. Sooranna, S. Ward, and M. Hancock. 2002. "Placenta as a link between amino acids, insulin-IGF axis, and low birth weight: evidence from twin studies." *Journal of Clinical Endocrinology and Metabolism* 87:308-15.

Bale, Judith R., Barbara J. Stoll, Adetokunbo O. Lucas, and Institute of Medicine (U.S.). Committee on Improving Birth Outcomes. 2003. *Improving birth outcomes : meeting the challenges in the developing world*. Washington, DC: National Academies Press.

Barker, D. J., P. D. Gluckman, K. M. Godfrey, J. E. Harding, J. A. Owens, and J. S. Robinson. 1993. "Fetal nutrition and cardiovascular disease in adult life." *Lancet* 341:938-41.

Becker, G. and N. Tomes. 1979. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy* 87:1153-89.

Behrman, J., R. Pollak, and P. Taubman. 1982. "Parental Preferences and Provision for Progeny." *Journal of Political Economy* 90:52-73.

Behrman, J. and M. Rosenzweig. 2004. "Returns to Birthweight." *The Review of Economics and Statistics* 86:586-601.

Behrman, J., M. Rosenzweig, and P. Taubman. 1994. "Endowments and the Allocation of Schooling in the Family and in the Marriage Market: The Twins Experiment." *Journal of Political Economy* 102:1131.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes. 2007. "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes*." *Quarterly Journal of Economics* 122:409-439.

Blickstein, I. 2005a. "Intrauterine Growth." Pp. 503-513 in *Multiple Pregnancy*, edited by I. Blickstein, Keith, L. Oxon: Taylor & Francis.

Blickstein, I. and R. B. Kalish. 2003. "Birthweight discordance in multiple pregnancy." *Twin Research* 6:526-531.

Blickstein, Isaac. 2004. "Is it normal for multiples to be smaller than singletons?" *Best Practice & Research Clinical Obstetrics & Gynaecology* 18:613-623.

Blickstein, Isaac. 2005b. "Growth Aberration in Multiple Pregnancy." *Obstetrics and Gynecology Clinics of North America* 32:39-54.

Boardman, J. D., D. A. Powers, Y. C. Padilla, and R. A. Hummer. 2002. "Low birth weight, social factors, and developmental outcomes among children in the United States." *Demography* 39:353-68.

Brooks, A. M., R. S. Byrd, M. Weitzman, P. Auinger, and J. T. McBride. 2001. "Impact of low birth weight on early childhood asthma in the United States." *Arch Pediatr Adolesc Med* 155:401-6.

Case, A., A. Fertig, and C. Paxson. 2005. "The lasting impact of childhood health and circumstance." *J Health Econ* 24:365-89.

Cleary-Goldman, J. and M. E. D'Alton. 2008. "Growth abnormalities and multiple gestations." *Semin Perinatol* 32:206-12.

Conley, D., K. W. Strully, and N. G. Bennett. 2006. "Twin differences in birth weight: the effects of genotype and prenatal environment on neonatal and post-neonatal mortality." *Econ Hum Biol* 4:151-83.

Currie, J. and R. Hyson. 1999. "Is the Impact of Health Shocks Cushioned by Socioeconomic Status? The Case of Low Birthweight." *American Economic Review* 89:245-250.

Currie, Janet. and Duncan Thomas. 2001. "Early Test Scores, Socioeconomic Status and Future Outcomes." *Research in Labor Economics* 20:103-132.

Datar, A. and A. Jacknowitz. 2009. "Birth Weight Effects on Children's Mental, Motor, and Physical Development: Evidence from Twins Data." *Maternal and Child Health Journal* 13:780-794.

Datar, A., M. R. Kilburn, and D. S. Loughran. 2010. "Endowments and parental investments in infancy and early childhood." *Demography* 47:145-62.

de Bie, H. M., K. J. Oostrom, and H. A. Delemarre-van de Waal. 2010. "Brain development, intelligence and cognitive outcome in children born small for gestational age." *Horm Res Paediatr* 73:6-14.

De Paepe, M. E., S. Shapiro, L. Young, and F. I. Luks. 2010. "Placental characteristics of selective birth weight discordance in diamniotic-monochorionic twin gestations." *Placenta* 31:380-6.

Fellman, J. and A. W. Eriksson. 2006. "Weinberg's differential rule reconsidered." *Hum Biol* 78:253-75.

Griliches, Z. 1979. "Sibling Models and Data in Economics: Beginnings of a Survey." *Journal of Political Economy* 87:S37-64.

Gringras, P. and W. Chen. 2001. "Mechanisms for differences in monozygous twins." *Early Human Development* 64:105-17.

Hack, K. E., J. B. Derks, S. G. Elias, A. Franx, E. J. Roos, S. K. Voerman, C. L. Bode, C. Koopman-Esseboom, and G. H. Visser. 2008. "Increased perinatal mortality and morbidity in monochorionic versus dichorionic twin pregnancies: clinical implications of a large Dutch cohort study." *BJOG* 115:58-67.

Hales, C. N., D. J. Barker, P. M. Clark, L. J. Cox, C. Fall, C. Osmond, and P. D. Winter. 1991. "Fetal and infant growth and impaired glucose tolerance at age 64." *BMJ* 303:1019-22.

Hall, J. G. 1996. "Twinning: mechanisms and genetic implications." *Curr Opin Genet Dev* 6:343-7.

Hall, J. G. 2003. "Twinning." *Lancet* 362:735-43.

Hankins, G. V. and G. R. Saade. 2005. "Factors influencing twins and zygosity." *Paediatr Perinat Epidemiol* 19 Suppl 1:8-9.

Hanley, M. L., C. V. Ananth, S. Shen-Schwarz, J. C. Smulian, Y. L. Lai, and A. M. Vintzileos. 2002. "Placental cord insertion and birth weight discordancy in twin gestations." *Obstetrics and Gynecology* 99:477-82.

Hoekstra, C., Z. Z. Zhao, C. B. Lambalk, G. Willemsen, N. G. Martin, D. I. Boomsma, and G. W. Montgomery. 2008. "Dizygotic twinning." *Hum Reprod Update* 14:37-47.

Imaizumi, Y. 2003. "A comparative study of zygotic twinning and triplet rates in eight countries, 1972-1999." *J Biosoc Sci* 35:287-302.

Ivanovic, D. M., E. R. Llop, J. A. Alvear, H. T. Perez, N. S. Diaz, B. D. Leyton, A. F. Almagia, C. G. Larrain, J. E. Alvarez, Y. F. Herrera, and E. L. Hazbun. 2006. "Chilean school-age children twin registry: planning, sampling and implications." *Twin Res Hum Genet* 9:743-6.

Kramer, M. and C. Victora. 2001. "Low birth weight and perinatal mortality." Pp. 57–69 in *Nutrition and health in Developing countries*, edited by R. D. Semba and M. W. Bloem. Totowa, NJ.

Leuven, E. and B. Sianesi. 2003. "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing."

Machin, G. 2009. "Non-identical monozygotic twins, intermediate twin types, zygosity testing, and the non-random nature of monozygotic twinning: a review." *Am J Med Genet C Semin Med Genet* 151C:110-27.

Mardones, F., L. Villarroel, L. Karzulovic, S. Barja, P. Arnaiz, M. Taibo, and F. Mardones-Restat. 2008. "Association of perinatal factors and obesity in 6- to 8-year-old Chilean children." *Int J Epidemiol* 37:902-10.

Mathews, T. J. and M. F. MacDorman. 2008. "Infant mortality statistics from the 2005 period linked birth/infant death data set." *Natl Vital Stat Rep* 57:1-32.

Matte, T. D., M. Bresnahan, M. D. Begg, and E. Susser. 2001. "Influence of variation in birth weight within normal range and within sibships on IQ at age 7 years: cohort study." *BMJ* 323:310-4.

Nikkels, P. G., K. E. Hack, and M. J. van Gemert. 2008. "Pathology of twin placentas with special attention to monochorionic twin placentas." *J Clin Pathol* 61:1247-53.

Oreopoulos, P., M. Stabile, R. Walld, and L. L. Roos. 2008. "Short-, medium-, and long-term consequences of poor infant health - An analysis using siblings and twins." *Journal of Human Resources* 43:88-138.

Pallotto, E. K. and H. W. Kilbride. 2006. "Perinatal outcome and later implications of intrauterine growth restriction." *Clin Obstet Gynecol* 49:257-69.

Paneth, N. S. 1995. "The problem of low birth weight." *Future Child* 5:19-34.

Pitt, M., M. Rosenzweig, and M. Hassan. 1990. "Productivity, Health, and Inequality in the Intrahousehold Distribution of Food in Low-Income Countries." *American Economic Review* 80:1139.

Resnik, R. 2002. "Intrauterine growth restriction." *Obstet Gynecol* 99:490-6.

Rose, H. 2006. "Do gains in test scores explain labor market outcomes?" *Economics of Education Review* 25:430-446.

Rosenzweig, M. and T. Schultz. 1982. "Market Opportunities, Genetic Endowments, and Intrafamily Resource Distribution: Child Survival in Rural India." *American Economic Review* 72:803.

Rosenzweig, M. and K. Wolpin. 1988. "Heterogeneity, Intrafamily Distribution and Child Health." *The Journal of Human Resources* 23:437-61.

Rosenzweig, M. and J. Zhang. 2009. "Do Population Control Policies Induce More Human Capital Investment? Twins, Birth Weight and China's “One-Child” Policy." *The Review of Economic Studies* 76:1149-74.

Royer, H. 2009. "Separated at Girth: US Twin Estimates of the Effects of Birth Weight." *American Economic Journal: Applied Economics* 1:49-85.

Sakata, M., M. Utsu, and K. Maeda. 2006. "Fetal circulation and placental blood flow in monochorionic twins." *The Ultrasound Review of Obstetrics & Gynecology* 6:135-140.

Scott, L. 2002. "The origin of monozygotic twinning." *Reprod Biomed Online* 5:276-84.

Sherer, D. M. and M. Y. Divon. 1997. "Fetal growth in multifetal gestation." *Clin Obstet Gynecol* 40:764-70.

Strauss, R. S. 2000. "Adult functional outcome of those born small for gestational age: twenty-six-year follow-up of the 1970 British Birth Cohort." *JAMA* 283:625-32.

Victoria, A., G. Mora, and F. Arias. 2001. "Perinatal outcome, placental pathology, and severity of discordance in monochorionic and dichorionic twins." *Obstetrics and Gynecology* 97:310-5.

Weinberg, W. . 1902. "Beitrage zur physiologie und pathologie der mehrlingsgebuten beim menschen " *Archiv fur die Gesamte Psysiologie des Menschen und der Tiere* 88:346-430.